

Preface

Generation Challenges 2010 was the fourth round of shared-task evaluation competitions (STECs) that involve the generation of natural language; it followed the Pilot Attribute Selection for Generating Referring Expressions Challenge in 2007 (AS-GRE'07) and Referring Expression Generation Challenges in 2008 (REG'08), and Generation Challenges 2009 (GenChal'09). More information about all these NLG STEC activities can be found via the links on the Generation Challenges homepage (<http://www.nltg.brighton.ac.uk/research/genchal10>).

Generation Challenges 2010 brought together three sets of STECs: the three GREC Challenges, GREC Named Entity Generation (GREC-NEG), Named Entity Reference Detection (GREC-NER), and Named Entity Reference Regeneration (GREC-Full), organised by Anja Belz and Eric Kow; the Challenge on Generating Instructions in Virtual Environments (GIVE) organised by Donna Byron, Justine Cassell, Robert Dale, Alexander Koller, Johanna Moore, Jon Oberlander, and Kristina Striegnitz; and the new Question Generation (QG) tasks, organised by Vasile Rus, Brendan Wyse, Mihai Lintean, Svetlana Stoyanchev and Paul Piwek.

In the GIVE Challenge, participating teams developed systems which generate natural-language instructions to users navigating a virtual 3D environment and performing computer-game-like tasks. The seven participating systems were evaluated by measuring how quickly, accurately and efficiently users were able to perform tasks with a given system's instructions, as well as on subjective measures. Unlike the first GIVE Challenge, this year's challenge allowed users to move and turn freely in the virtual environment, rather than in discrete steps, making the NLG task much harder. The evaluation report for the GIVE Challenge can be found in this volume; the participants' reports will be made available on the GIVE website (<http://www.give-challenge.org/research>) at a later stage.

The GREC Tasks used the GREC-People corpus of introductory sections from Wikipedia articles on people. In GREC-NEG, the task was to select referring expressions for all mentions of all people in an article from given lists of alternatives (this was the same task as at GenChal'09). The GREC-NER task combines named-entity recognition and coreference resolution, restricted to people entities; the aim for participating systems is to identify all those types of mentions of people that are annotated in the GREC-People corpus. The aim for GREC-Full systems was to improve the referential clarity and fluency of input texts. Participants were free to do this in whichever way they chose. Participants were encouraged, though not required, to create systems which replace referring expressions as and where necessary to produce as clear and fluent a text as possible. This task could be viewed as combining the GREC-NER and GREC-NEG tasks.

The first Question Generation challenge consisted of three tasks: Task A required questions to be generated from paragraphs of texts; Task B required systems to generate questions from sentences, and Task C was an Open Task track in which any QG research involving evaluation could be submitted. At the time of going to press, the QG tasks are still running; this volume contains a preliminary report from the organisers.

In addition to the four shared tasks, Generation Challenges 2010 offered (i) an open submission track in which participants could submit any work involving the data from any of the shared tasks, while opting out of the competitive element, (ii)

an evaluation track, in which proposals for new evaluation methods for the shared task could be submitted, and (iii) a task proposal track in which proposals for new shared tasks could be submitted. We believe that these types of open-access tracks are important because they allow the wider research community to shape the focus and methodologies of STECs directly.

We received three submissions in the Task Proposals track: an outline proposal for tasks involving language generation under uncertainty (Lemon et al.); a proposal for a shared task on improving text written by non-native speakers (Dale and Kilgarriff); and a proposal for a surface realisation task (White et al.).

Once again, we successfully applied (with the help of support letters from many of last year's participants and other HLT colleagues) for funding from the Engineering and Physical Sciences Research Council (EPSRC), the main funding body for HLT in the UK. This support helped with all aspects of organising Generation Challenges 2010, and enabled us to create the new GREC-People corpus and to carry out extensive human evaluations, as well as to employ a dedicated research fellow (Eric Kow) to help with all aspects of Generation Challenges 2010.

Preparations are already underway for a fifth NLG shared-task evaluation event next year, Generation Challenges 2011, which is likely to include a further run of the GIVE Task, a second run of the QG Challenge, and a pilot surface realisation task. We expect that results will be presented at ENLG'11.

Just like our previous STECs, Generation Challenges 2010 would not have been possible without the contributions of many different people. We would like to thank the students of Oxford University, KCL, UCL, Brighton and Sussex Universities who participated in the evaluation experiments, as well as all other participants in our online data elicitation and evaluation exercises; the INLG'10 organisers, Ielka van der Sluis, John Kelleher and Brian MacNamee; the research support team at Brighton University and the EPSRC for help with obtaining funding; and last but not least, the participants in the shared tasks themselves.

July 2010

Anja Belz, Albert Gatt and Alexander Koller